Original Article

"Challenges Of Big Data Revolution In Health Care System"

Al Shammari, Naif Dhidan[1]*, Al Bedaiwey, Khalid Abdullah[2], Al Otaibi, Mansour Bakheet[3], Hadadi, Salha Mohammed[4], Al Otaibi, Ashwaq Turky[5], Al Otaibi, Fahad Marzouq[6]

[1]*Ministry Of National Guard Health Affairs, Email: Shammarinay@mngha.med.sa
[2]Ministry Of National Guard Health Affairs, Email: Albdawekh@mngha.med.sa
[3]Ministry Of National Guard Health Affairs, Email: Alotaibima20@mngha.med.sa
[4]Ministry Of National Guard Health Affairs, Email: Hadadisa@mngha.med.sa
[5]Ministry Of National Guard Health Affairs, Email: Alotaibias@mngha.med.sa
[6]Ministry of National Guard Health Affairs, Email: Alotaibifa14@mngha.med.sa

***Corresponding Author: - Al Shammari, Naif Dhidan**

[1]*Ministry Of National Guard Health Affairs, Email: Shammarinay@mngha.med.sa
DOI: 10.47750/pnr.2022.13.02.90

Abstract

The big data revolution and state-of-the-art data analytics are driving advances in healthcare, including earlier diagnosis, better care coordination, and better outcomes at lower costs. Technologies for storing, merging, and analyzing data have had far-reaching effects in business and medicine. As the healthcare sector begins to tap into the potential of big data, concerns around security and privacy have come to the fore. The widespread potential of the Internet of Medical Things (IoMT) has piqued the interest of academics. Sensitive healthcare data acquired by smart healthcare sensors and IoT-enabled medical equipment is sent to server nodes in a secure manner, with no human intervention required. It is difficult to safely collect and transfer healthcare data towards Fog and cloud servers for a number of reasons, including but not limited to concerns about security and privacy. In this study, we looked at the available surveys to determine the privacy and security threats brought on by the big data revolution in the healthcare sector. We have conducted a survey and analytical study of fog-assisted secure healthcare data collecting techniques.

Keywords: Healthcare; big data security; privacy; security analytics, IoT, healthcare, smart medical devices, fog computing, data aggregation.

1. INTRODUCTION

The term "big data" is used to describe data sets that are too massive or complicated to be analyzed using conventional techniques. Machine learning is used instead, which consists of self-updating algorithms that construct prediction models by identifying patterns in data. The promise of a "big data revolution" in healthcare has been made so often in recent years, it has prompted scholars to wonder why it has not materialized. There is a deeper problem that has been appropriately highlighted, despite the fact that certain technological impediments have been discovered: many of the data are of low quality and in the form of tiny, incompatible datasets (De at al., 2016).

It is challenging to deploy machine learning on a wide scale in healthcare due to the current processes surrounding data gathering, curation, and sharing. We need to create, analyze, and apply cutting-edge health data standards to assure data quality, make it possible to pool information from diverse institutions, and provide researchers and others with fast access to datasets. These conditions are still lacking before machine learning can be implemented. The hoopla around machine learning sometimes overshadows the fact that it is still merely a tool for data science, one that has its own specific set of prerequisites and limits. All healthcare technologies have to fit inside a wide variety of human limits, from the molecular to the social and political, which is something the hype doesn't take into account. When combined with the intricate infrastructure of clinical practice and healthcare delivery, each of these factors will impede progress (Najafabadi et al., 2015).

The Internet of Things (IoT) is a network of interconnected electronic gadgets that can collect and relay data. The Internet of Things (IoT) relies on a wide range of applications, including but not limited to healthcare, mining, buildings, cities, agriculture, transportation, industries, and automated systems, all of which benefit greatly from the proliferation of intelligent sensory elements and wearable smart gadgets. Smart medical gadgets provide communication between patients and other connected equipment, streamlining healthcare processes. The IoMT is quickly developing into a crucial part of the healthcare system. It collects data in a variety of formats and sends it to remote servers so that it may be used to improve healthcare delivery. IoMT is the glue that makes smart healthcare work. Therefore, a sustainable answer is necessary to address the shortcomings of current approaches to IoT-based smart healthcare. In order to enhance the effectiveness and efficiency of medical treatment, medical equipment allow for remote monitoring of patients. Sensors worn by patients provide data about their health to neighboring computers (Mahmud et al., 2018).

The number of hospital IT security breaches has been rising steadily over the previous decade. After the theft of an unencrypted USB flash drive holding patient records in 2013, Kaiser Permanente (one of the largest non-profit healthcare organizations in the US) alerted its 49,000 patients that their health information had been stolen (McCann, 2013). Verizon's forensic investigation and security team analyzed data from 47,000 reported security events and confirmed 621 data breaches in 2012, according to the company's data breach investigation report. Additionally, 94% of hospitals suffered at least one security breach in the last two years, according to a survey on patient privacy and data protection (Ponemon, 2012).

Rather of coming from outside sources, most attacks came from within. The analysis also revealed that the United States, China, and Eastern Europe were the origins of the assaults. (Romania recording the highest number of external attacks). Increases in security breaches are anticipated because of the dynamic nature of the risk environment and the emergence of novel threats and vulnerabilities. More people signing up for health insurance as a result of the Affordable Care Act will be a prime target for cybercriminals, unleashing a wave of healthcare breaches over the next several years. Violating the Health Insurance Portability and Accountability Act (HIPAA) or the Health Information Technology for Economic and Clinical Health Act (HITECH) in the United States is illegal if a breach of electronic health record security compromises patient privacy (Hanson et al., 2009). Therefore, protecting the confidentiality of electronic health records is crucial.

2. LITERATURE REVIEW

2.1. Big data revolution

Information (data) about customers' wants, requirements, habits, and preferences has grown at an exponential rate thanks to the advent of big data. Information is abundant now because to developments in computers, communications, and the Internet; devices, applications, and online services provide fresh perspectives on our identities as individuals, communities, customers, and patients (Chen et al., 2012). Everyday life involves being exposed to a plethora of transactions, many of which generate data. Everything we do—from the food we eat to the clothing we buy to the movies we watch to the places we reside to the medical professionals we see to the social services we use—are all recorded transactions. Furthermore, models may be constructed to aid in future prediction by associating and correlating data on trends, wants, and seemingly unrelated factors.

With the advent of big data, we now have access to a flood of data that provides a detailed portrait of the human condition. Innovations in illness detection, care coordination, better health outcomes, and cost reductions are being fueled by the big data revolution and the sophisticated data analytics it enables (Bates et al., 2014). The interconnectedness of health, the environment, and socioeconomic forces has also been brought to light by cutting-edge analytical tools such as machine learning (for instance, sifting through massive amounts of data to develop a tool to predict chronic homelessness) and artificial intelligence (for instance, using genomic data to better diagnose breast cancer) (Toros & Flaming, 2017).

Supply chain and marketing effectiveness as well as revenue growth have benefited from improvements in data processing, analysis, and visualization that have led to actionable insight. When compared to its rivals, Amazon.com is able to link customers with items more quickly, intelligently, and affordably because to the use of big data (Markman, 2017). The big data revolution will not only improve the efficiency and effectiveness of connecting businesses and consumers, but also individuals and community organizations.

2.2. Need of Big data Revolution in Healthcare

The huge gap between healthcare spending and results highlights the critical importance of making healthcare more efficient. In 2017, national health expenditure increased by 3.9% to $3.5 trillion, or 17.9% of GDP, according to the U.S. Centers for Medicare & Medicaid Services. Both Medicare (at $705.9 billion) and Medicaid (at $581.9 billion) are included in this total (Center for Medicare and Medicaid Services, 2017). In the midst of escalating healthcare expenditure and poor health outcomes, the Institute for Healthcare Improvement (IHI) came up with the term the "Triple Aim." This word articulates the need to enhance the patient experience, boost population health, and lower healthcare costs per person. In order to attain the elusive "triple aim" of lowering healthcare costs, improving quality, and expanding access, healthcare providers are beginning to incorporate the insights gained from business analytics (Bates et al., 2014).

The healthcare system is a massive dynamic system with several intricate components. A person's health is improved through a network that includes chronic diseases, primary care physicians, specialists, clinics, pharmacies, and emergency services. Data may be collected and evaluated from all interactions and subsystems to inform the design of more precise treatments. Data analytics, such as data mining, machine learning, and predictive modeling, are useful for pinpointing groups likely to have complicated demands and incur high healthcare costs. Furthermore, they are stressing the value of researching the underlying socioeconomic causes of disease. Opportunities to reduce costs and enhance quality of care may be found in a wide range of data management systems now available to hospitals and healthcare providers (Roski et al., 2014).

The application of machine learning to improve the prediction of unfavorable birth risk among pregnant women was investigated, for instance, by Pan et al. (2017). Results showed that machine learning algorithms performed 36% better than risk-based assessments, leading them to the conclusion that this method of assessing risk makes for more efficient use of scarce resources (Pan et al., 2017). Better results and lower costs can be achieved by identifying those most in need of care as early as possible. Business analytics that are too complicated for most social workers to employ are already being put to use in the healthcare industry.

2.3. Challenges of Big data's revolution in healthcare

Managing patient care and medical innovation has become more important than ever as the healthcare industry continues to grow. Machine learning algorithms may quickly provide forecasts and suggestions for the patient and his attending Doctor or Physician by finding statistical connections in a massive global medical data set. The most fundamental approaches to use big data in healthcare provide registries of medical data from which we may all benefit. With this data in hand, we may foresee future "waves" of sickness (Carvalho et al., 2017).

2.3.1. Prediction of disease development: Using these digital health records, physicians have found links between previously unrelated ailments (Patterson, 2016). The medical practitioner may now anticipate whether or not a diabetic patient will acquire dementia using the risk assessment approach created in 2013. The U.S. military is attempting to minimize the rate of veteran suicide by employing the same approach.

2.3.2. Identification of genetic markers in oncology: After examining the most prevalent cancers, researchers at UCT came to the conclusion that each form of cancer is caused by a unique set of genetic mutations. Cancers of the breast, intestine, lungs, ovaries, and brain have been found to have their own unique genetic signatures. The study's coordinator argues that the research group couldn't have made the finding without the availability of massive amounts of data.

2.3.3. Predicting the health of babies: The Toronto Children's Hospital has launched the Artemis initiative. Real-time data on neonates is collected and analyzed by the hospital information system. Every one of a child's 1260 indications may be monitored by the system. As a result, pediatricians and nurses may better prepare for the prevention of illnesses by predicting the individual children's susceptibility to them.

2.3.4. Prediction of risk factors in surgery: QPID (qpid.apache.org) is a tool used by MGH doctors. Apache QPID is an open-source messaging system that uses the Advanced Message Queuing Protocol (AMQP) analytical system to keep track of vital patient data as it is collected and processed. The QPID method is also useful in healthcare for predicting surgical risk (Kalejahi et al., 2019). Treatment procedures are automatically searched for by the QPID system, and the findings are shown on a screen along with a determined red, yellow, or green risk signal.

2.3.5. Creating new drugs: The field of drug modeling stands to benefit most from the use of Big Data. Semantic Hub, an IT startup, is now evaluating and developing the potential for producing new medicinal treatments with the use of data obtained from the medical profession.

2.3.6. Improving the quality of clinical trials: Companies may improve the efficiency of clinical trials by using Big Data technology. Analytical tools may now choose patients who are most likely to pass a drug test by drawing on patient information from many databases (Yang et al., 2016). Researchers can now keep tabs on study participants in real time because to technological advances in telemedicine, gadgets, and wearable devices.

2.3.7. Identification of drug side effects: Before beginning clinical trials, adverse effects of individual substances and components can be predicted using big data. Companies can save money and lives by evaluating hundreds of medication properties using an analytical approach before releasing the drug to the public. Traditional randomized clinical trials remain the foundation of drug testing today, but there has been a recent uptick in interest in data collected on medication usage in everyday life. According to Reuters, in 2017 there were more than 300 experiments involving the analysis of large datasets. Why does it seem like such a huge issue to deal with Big Data? To claim that big data holds the key to any and all relevant information would be incorrect. There are several obstacles to overcome when trying to get useful information from a large dataset.

2.3.8. Unstructured data: If the information you need is in text format, a sophisticated search algorithm can help you find it. The issue occurs when the data is presented in an unreadable manner, such as a video or a voice. The technology exists to transcribe audio or video into text, but the resulting data volumes will be incredibly vast, creating storage and sorting challenges. It will be very expensive to filter and analyze the about 78% of medical data that is not in a usable format. I. Serious potential for data distortion. Some skeptics go so far as to call Big Data a massive fraud. Big data was slammed by a wave of anger after Google Flu Trends was widely panned as a failure. To give just one example, Google's initiative failed to detect and report the US pandemic of 2013. After two years of investigation, researchers from many American colleges concluded that the analysis had been producing more frequently inaccurate findings. Healthcare information systems frameworks involving large amounts of data.

2.4. Data Security and Privacy Issues in Healthcare Using Big Data Revolution

Concerns about patients' security and privacy are amplified by the widespread adoption of big data in healthcare. In the beginning, patient data is housed in data centers that offer varied degrees of protection. Furthermore, most healthcare data centers are HIPAA-compliant, however this does not ensure the security of patient records. HIPAA's emphasis on guaranteeing security policies and procedures rather than on actually putting them into practice is the cause for this. In addition, the storage, processing, and communication infrastructures are being stressed by the influx of enormous data sets from many sources.

Large data sets are intrinsically heterogeneous, making it impossible to apply traditional security measures directly to them. As healthcare cloud solutions gain in popularity, the challenge of securing large-scale distributed SaaS systems becomes more complicated due to the wide variety of data sources and formats involved. Therefore, big data governance is required before data can be used for analytics.

2.4.1. Data governance

The first step in controlling and managing healthcare data will be data governance as the healthcare industry shifts toward a value-based business model using healthcare analytics. The objective is to create a unified data representation that takes

into account both global norms (such as LOINC, ICD, SNOMED, CPT, etc.) and regional and national variations. BSN data is currently heterogeneous and would need normalization, standardization, and governance before analysis can begin.

2.4.2. Real-time security analytics

The ever-expanding healthcare sector has made real-time risk analysis and threat source prediction imperative. The healthcare sector is currently under attack from a wide variety of advanced threats, including Distributed Denial of Service (DDoS) and stealthy malware. Furthermore, social engineering assaults are on the rise, and it is challenging to forecast the dangers associated with these attacks without taking human cognitive behavior into account. For older individuals in particular, cognitive bias might be a problem. The term "cognitive bias" refers to an error in reasoning that can lead to irrational conclusions about other persons and circumstances. (Marshall et al., 2013). A man-in-the-middle attack, for instance, may be carried out by convincing an old patient to accept a bogus digital X.509 certificate. An end-to-end authentication system needs to be planned with these kinds of cases in mind.

Implementing security in resource-constrained networks has proved difficult in the IoT context and will continue to get more complicated as the number of IoT devices grows. Common symmetric and asymmetric key distribution and revocation procedures, for instance, will break down when used to a network of a billion IoT gadgets. So, for IoT to integrate large data in a cloud setting, we need new scalable key management systems that result in smooth inter-operability between different networks (such IoT and conventional IP networks). Any cloud-based SaaS solution that stores Protected Health Information must be built around security analytics as the healthcare sector increasingly relies on new big data technologies for more informed decision making. (PHI). In addition, novel approaches to risk management will be guided by real-time intelligence in the realm of security. This allows healthcare IT companies to keep an eye on potential threats in real time and take precautions before they have an impact on the healthcare industry (Patil and Chen, 2013).

2.4.3. Privacy-preserving analytics

The ever-expanding healthcare sector has an urgent need for continuous risk analysis and the identification of potential causes of disruption. Distributed denial of service (DDoS) and other forms of sophisticated assault, as well as more covert malware, are already flooding the healthcare sector. In addition, incidents of social engineering are on the rise, and the hazards connected with them are hard to foresee without taking into account individuals' underlying psychological make-up. For example, cognitive bias may have a role, especially when dealing with older people. To err in one's judgment in such a way as to make incorrect conclusions about other individuals or events is an example of "cognitive bias" (Marshall et al., 2013). A man-in-the-middle attack can be launched, for instance, if a senile patient is convinced to accept a fraudulent digital X.509 certificate. In creating an end-to-end authentication system, it is important to consider such use cases.

Security in the Internet of Things has proven difficult to implement in resource-limited networks, and it will only become more complicated as the number of IoT devices proliferates. Key distribution and revocation systems that have been used in the past, both symmetric and asymmetric, cannot be scaled up to a billion IoT devices. For this reason, IoT's integration of large data in a cloud environment requires novel scalable key management solutions leading to smooth inter-operability between different networks (such as IoT and conventional IP networks). Whenever a healthcare provider uses a software as a service (SaaS) platform to store patients' protected health information in the cloud, security analytics should be at the center of the solution's architecture. (PHI). Also, novel approaches to risk management will be guided by real-time security intelligence. Providers of healthcare IT are therefore in a position to keep tabs on potential threats in real time and head them off at the pass (Patil and Chen, 2013).

3. METHODOLOGY

Aggregated data that is both secure and private is essential at both the end node device and the fog node. Since edge devices collect data from sensor nodes and transfer it to the cloud server, authenticating these devices is a crucial duty for keeping data secure. Two distinct varieties of cryptography are used for data-aggregation security. First, with the help of public and private keys for encryption and decryption, asymmetric cryptography allows for safe data aggregation. Second, symmetric cryptography allows for the safe collection of data with just one key needed for encryption and decryption. The confidentiality of patients' medical records is crucially dependent on privacy protection and encryption-based security (Hou et al, 2019). Compressing data has played an important part in healthcare data collecting with security and privacy. The compression ratio determines how much space may be saved when data is compressed. It lessens the time it takes for data to go from the edge node to the cloud and the amount of money spent on communication and computing. Data aggregation in smart healthcare and device interaction scenarios offered by the fog is shown in Figure 1.

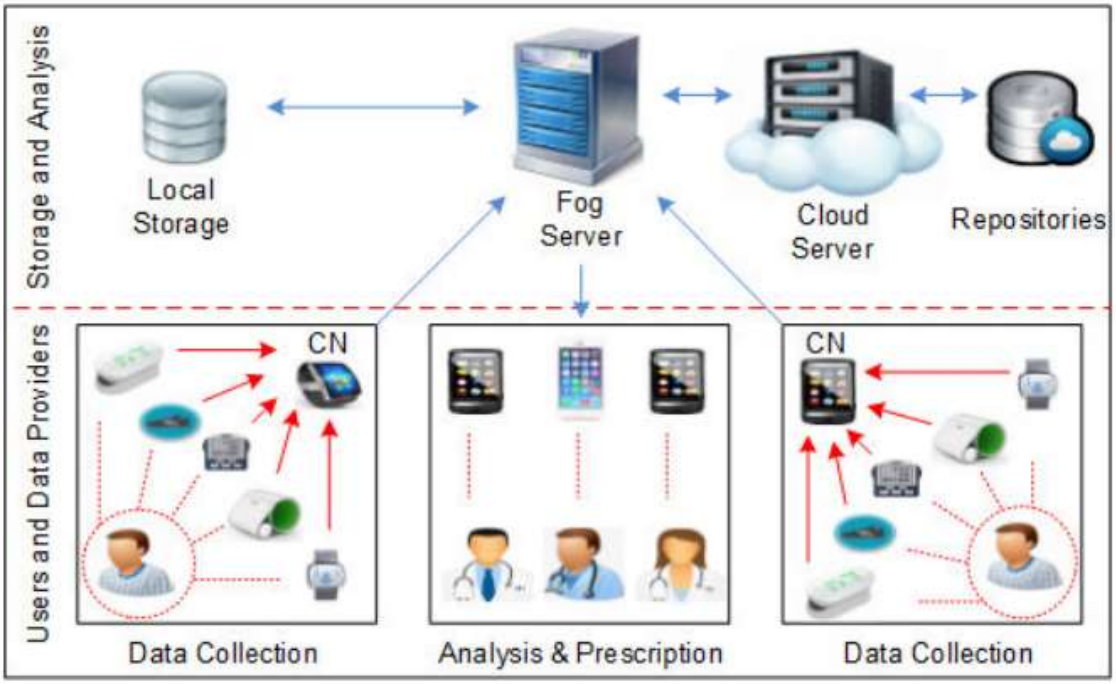

Figure 1. Collecting Data via smart Medical Devices

The patient has many sensors connected to their person in order to gather health information. Smartphones, tablets, and wearable smart gadgets are examples of sensing devices that convey this information to the data aggregator node. Sensitive health records are sent to the Fog server from smart collection nodes. Edge nodes provide local compression and authentication on patient data before storing it. A fog node transforms information into a cloud-friendly format. A cloud server retrieves information from a fog node and saves it to a cloud database. Non-delay-tolerant data is sent to the cloud with priority by the fog server, which also handles real-time healthcare data. Data that has been compiled and uploaded to a cloud repository is compressed before being stored there. Medical practitioners and other authorized users can access their patient records using cloud-based storage systems. In order to obtain the necessary data, the edge node must first receive a request from an authorized user. If the desired device is within range of an edge node, the node will transmit the required data to it. Without it, edge devices access the cloud for the necessary information.

4. RESULTS AND DISCUSSION

The security of the Internet of Things (IoT) may be broken down into three distinct categories from a data scientist's point of view: one-stop, multi-stop, and endstop. In the one-stop view, data is compiled at a single endpoint node before being sent and received over the internet. For this reason, secure and reliable data transmission over the internet necessitates lightweight cryptography. The multi-stop feature ensures that different sets of devices are always online or linked to a private network. As a result, there is a pressing demand for encrypted communication between several nodes. The applications are grounded in several domains, such as smart homes and smart hospitals (Gope et al., 2015). Challenges in privacy, forensics, and societal or legal norms arise from these factors, making it difficult for researchers to find workable solutions.

A methodology for both data aggregation from end nodes and the anonymity of those nodes is presented by the Anonymous and Secure Aggregation Scheme (ASAS). To protect user privacy, it uses homomorphic encryption and pseudonyms. End node devices anonymously communicate data to the fog node and provide assistance to the end node data via the cloud server, all while maintaining data integrity. The provided approach reduces the amount of data sent between the Fog and the cloud (Wang et al., 2018). Combining Homomorphic Paillier encryption with the Chinese Remainder Theorem yields a lightweight privacy-preserving data aggregation (LPDA) Scheme. Additionally, data is consolidated in one location.

Similarly, the risk of bad data insertion is mitigated with the use of a one-way hash chain approach. To immediately collect device Di individual data, edge nodes report aggregation based on Cis and receive the total number of c1s, c2s,..., cNs from all IoT devices in time slot TS. In order to perform the data aggregation operation illustrated by equation, fog devices use the secret key SN+1 to generate the hash H (TS) n.SN+1. (1). In this setup, Fog devices execute data processing locally before uploading filtered data to a cloud storage service and further subdividing devices based on their sensing capabilities. In terms of efficiency, it reduces the time and energy spent on calculation and communication.

$$\begin{cases} C_s = \left(\prod_{i=1}^{N} C_{is}\right).H\,(T_S)^{n.S_{N+1}} \\ mac_s = h\,(C_s\,||TS||\,sk) \end{cases} \quad (1)$$

The Anonymity and Privacy Preserving Aggregation (APPA) system does just that. In addition to protecting the authenticity of data collected from smart devices, it also allows for the anonymous updating of certificates. At the system's periphery, smart sensing devices (SDs) collect data and send it to the Fog nodes (FNs) over a communications network. It is a buffer between the public cloud server PCS and the data it receives, storing the data locally until it can be processed locally according to the cloud server's preferred format. Equation 2 shows the results of data collection by smart devices at Fog d1, d2,..., dn at time t T. Therefore, SDi is the i th smart device, and it selects a random number from rs Z N, Ci is made up of SD0 i s pseudonym, and it computes the message digest of encrypted data starting at position (i). Using the hash of the received ciphertext (H3 (Ci) mod n), SDi sends a data packet to the FNK node (Eq. 3).

$$\begin{aligned} C_i &= (Pseu_{SD_i})^{d_i}.r_s^n mod\ n^2 \\ &= (g^{r_i r_j})^{d_i}.r_s^n mod\ n^2 \quad (2) \end{aligned}$$

$$SD_i \rightarrow FN_K : \{C_i\,||\sigma_i||\,Crep_{SD_i}||TS\} \quad (3)$$

Data is aggregated at the Fog node, FNK, to check if (H3 (Ci)) mod n is true in equations (4) and (5). If the condition is met, the SDs begin sharing aggregated data using a pseudonym certificate. A report packet is generated by computing the message digest using the function CaH3 (Ca) mod n, and then transmitted to a public cloud server. The word rk is a random number chosen during pseudonym generation; (CrepSDi) is a pseudonym certificate at smart devices; and TS is a timestamp.

$$\begin{aligned} \sum_{i=1}^{n}(C_i.Crep_{SD_i}) &= \sum_{i=1}^{n}[\left(g^{r_i r_j}\right)^{d_i}.g^{r_j' r_z}]mod\ n^2 \\ &= g^{(d_1,d_2\ldots,d_n)r_j r_z} mod\ n^2 \quad (4) \end{aligned}$$

$$\begin{aligned} C_a &= \sum_{i=1}^{n}\left(C_i.Crep_{SD_i}\right).Crep_{FN_j}.g^{r_k'} \\ &= g^{(d_1,d_2\ldots,d_n)r_j r_z}.g^{r_i' r_z' r_k}.g^{r_k'} mod\ n^2 \\ &= g^{\sum_{1=1}^{n} d_i} mod\ n^2 \quad (5) \end{aligned}$$

Data received by PCS is analyzed computationally in the context of the desired format. Data may be kept safe and sent securely thanks to trusted third-party certification authorities (TCAs) and local certification authorities (LCAs).

$$FN_K \rightarrow PCS : \{\sigma_{C_a}||C_a||TS||Crep_{SD_i}\} \quad (6)$$

Nonetheless, this allows for sophisticated smart-device authentication on several levels. However, it is a practical option for performance when working with constrained hardware. The efficiency of the APPA method degrades when the number of connected devices grows big. A scalable and effective solution is needed to address this problem. It necessitates a high-performance, real-time healthcare data transfer system. An EoT security architecture for use in smart healthcare. Fully homomorphic encryption (FHE) is used to keep sensitive information safe. For real-time computing and local storage on the EoT devices, clustering-based approaches evaluate large-scale heterogeneous data.

Numerous gadgets produce data, which they send via cloud server. To lessen the burden on cloud servers, Internet of Things (IoT) devices sit between the cloud and end node devices. The two BGV keys, secretkey and publickey, are used in key generation. When given the plaintext m, homomorphic encryption produces the ciphertext c, where op is the number of computations. The homomorphic characteristics of BGV are given by the equation 7. Data that has been encrypted with FHE can be stored and analyzed. In this study, we employ clustering-based strategies for edge node local processing, namely the K-means clustering (KMC) algorithm and the Fuzzy C-Mean clustering (FCMC) algorithm.

$$m_1 op\ m_2 = DEC\,(Enc\,(m_1)\,op\,Enc\,(m_2)) \quad \forall m_1,\ m_2 \epsilon A_p \quad (7)$$

Data authentication and access control at wearable devices in the time-aware and space-aware scenario is provided by a cooperative privacy protected system. In a location-aware setting, the edge node uses MinHash authentication to protect sensitive data while also determining which patients' records are most similar to one another. Using cyphertext base encryption with bloom filters in a time-aware cloud server environment enables access control and yields an efficient data

structure. Issues including mutual authentication, privacy preservation, and data integrity preservation are elaborated and resolved in this smart healthcare architecture for edge and cloud-based hybrid computing. GNY logic underwent a security examination to verify its efficacy as a design.
Updated ciphertext provided by a (ABE/ABS) ensures data security in Fog computing. Additionally, use attribute-based signatures and ciphertext policy attribute-based encryption (CP-ABE). (ABS). Using ciphertext updates and computation outsourcing, it allowed for protected data access. It demonstrates multiple-policy, attribute-based data encryption. In order to encrypt and decrypt data, terminal nodes send ciphertext through the Fog node. To transmit data across cloud repositories, the data must be signed at the Fog node. Only the recipient can decode the ciphertext if their characteristics match those specified by the access regulations. It also allows for restricted access to data and secure ciphertext updates.

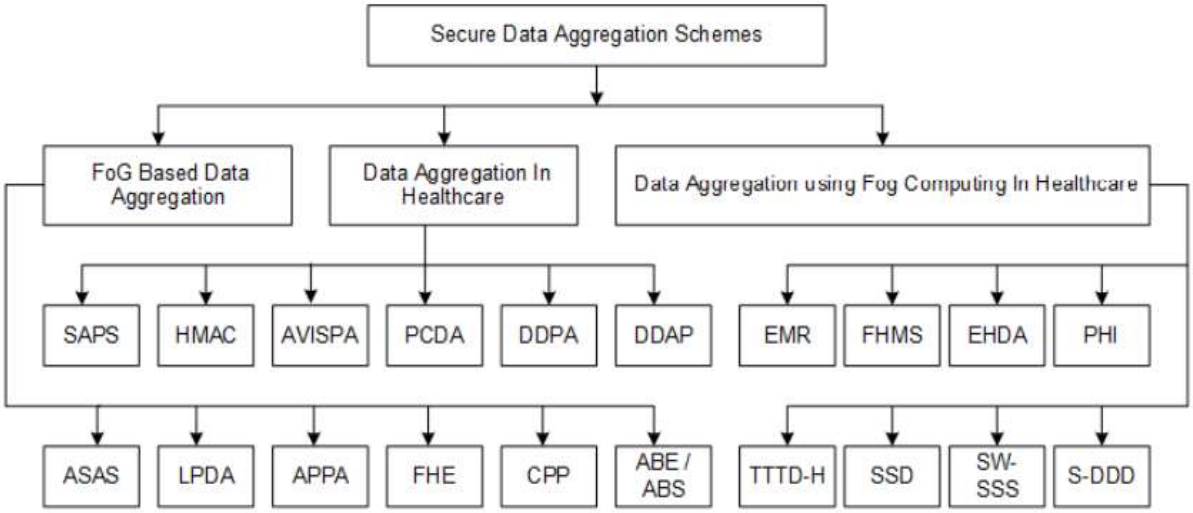

Figure 2. Classification of safe methods for compiling data

As big data revolutionizes healthcare, it's more important than ever to ensure the safety and privacy of patients using these new systems. Hosting businesses will be less willing to offer enormous healthcare data for centralized processing as healthcare clouds with big data gain popularity. As a result, we anticipate utilizing distributed processing across many clouds and relying on group wisdom. As healthcare clouds combine and integrate massive volumes of data from different networks, secure patient data management is inescapable. Proactive health care and wellness will also be propelled by real-time analytics that are both safe and protect users' privacy. In this study, we examine some of the healthcare industry's security and privacy challenges, and we predict that advances in computing, storage, and communication will be required to keep up with the increased need for protecting sensitive patient information.
However, there are still numerous challenges to overcome, such as the development of low-cost sensing devices, sophisticated algorithms for lifelogging data, privacy, and security, before IoT-enabled healthcare solutions can fully realize their potential. Many obstacles to developing effective and secure data aggregation systems for fog-assisted IoMT are discussed in this section. The Healthcare industry faces a number of complex problems. However, we only focus on the ones that matter. Overcoming these obstacles is crucial for the future. As a result of the COVID-19, there is a pressing demand for safe patient data collecting and remote monitoring. Our study identifies critical issues within the healthcare sector, therefore pointing the way for future research towards IoT-enabled healthcare.
Protecting the data and infrastructure of an IoMT is a top technological priority. Safer networks and hardware allow for more reliable IoMT. If a device meets the trust management criteria, it will be safe to use in conjunction with other devices. Data exchange between sensors and servers on open networks requires a high level of security. Avoiding criminals who steal sensitive health data or conduct covert physical attacks on open networks is challenging. When data is aggregated, many parties have access to a patient's private data, creating a security risk. Designing security solutions for MWSNs is difficult for a number of reasons. It is difficult to devise a foolproof defense against all potential security breaches (Yang et al., 2017).
Protecting sensitive data during transmission to a Fog or cloud server is a major concern, as is the ability to identify and respond to many threats. Protecting and detecting malware at the Fog node in the healthcare sector is an open area of study. Protecting fog-assisted systems against malware assaults will need the development of acceptable algorithms that offer continuous protection and resource calculation. There are many edge devices in a fog-based system, and any one of them can be compromised and used to launch a denial-of-service attack. Because smartphones and other smart gadgets are so susceptible to malware infiltration, they are often used as fog nodes. Therefore, it is necessary to develop AI techniques or deploy malware detection devices at edge nodes to prevent these security breaches. In order to corrupt data or impede the center layer's communication with the outside layers, attackers focus on the intermediate layer.
Edge computing devices provide a significant challenge to security research. WBAN encrypts private information alongside the ciphertext to keep data storage and other costs to a minimum. Making a safe system that uses less healthcare resources is difficult, though. There are a number of obstacles that the Internet of Things must overcome in order to

aggregate data from smart devices. Changes can be made to the patient's private medical records in this scenario. As a result, fog nodes are vulnerable to security risks that don't affect cloud computing (Lin et al., 2017).
Quality of service (QoS) is the assignment of various priority to different applications in order to guarantee a minimum acceptable level of data transmission performance. When shifting work from a busier node to a less busy one in a Fog-based healthcare paradigm, maintaining service quality is paramount. The major difficult challenge is protecting sensitive patient data as it travels from the end node to the server node (Paul et al., 2018). Latency refers to the amount of time it takes for a data packet to travel between two node devices in an IoT network. The time it takes for data to be sent to and received from a device is known as its latency, or round trip time. Latency refers to the time it takes for anything to happen, such when one system is waiting for another to finish processing. The latency and bandwidth are two performance factors that affect the quality of service. In other words, QoS entails minimal overhead for the Fog node in terms of processing time and energy. In a healthcare system built on the Fog, Quality of Service (QoS) is a major obstacle.
Connecting various technologies, smart devices, and apps is what makes the Internet of Things so vital to enhancing human well-being. Medical facilities and their physical roles are taken into account by healthcare plans. In this scenario, intelligent medical devices collect patient data and send it to central server nodes. It is difficult to send data to the Fog/cloud server in a safe and effective manner.
While "big data" showed initial promise in healthcare, the unique restrictions of clinical science, including data quality, privacy, and regulatory laws, made its first promises difficult to realize. We spoke about these ideas because we want to create a comprehensive strategy for getting health research results directly into patient care. We suggest that current big data systems are in their infancy, and that health care big data may not realize its full potential until these basic difficulties are addressed. We conclude that in order to advance, a broader group of institutions must be ready to invest in technology for de-identifying private patient data so that it may be shared extensively for scientific study, and that this data must be full, exact, and time stamped. The scientific and regulatory communities, in light of the global quality of care gap, need to develop novel approaches to understanding causal relationship from data captured during routine health care, which would complement existing gold standard methods like randomized controlled trials (RCTs) and identify the connection between clinical practice and outcomes.

5. CONCLUSION

As we can see, many different security methods have been established and may be employed in the healthcare industry to get unauthorized access to private patient information. There are a lot of variables that can affect how well these security measures function, including the nature of the workplace, the size of the business, and the available money. However, it is challenging to determine when an e-healthcare system's information and the environment in which it is managed are safe from unauthorized access. It is well knowledge that complete safety cannot be guaranteed. As big data revolutionizes healthcare, it's more important than ever to ensure the safety and privacy of patients using these new systems. As big data healthcare clouds gain popularity, hosting businesses will be less willing to allow the transfer of large amounts of healthcare data for analysis. This is why we imagine a world where data is processed in parallel across several clouds, drawing on the collective wisdom of the internet. As healthcare clouds combine and integrate massive volumes of data from various networks, secure patient data management is necessary. Proactive health care and wellbeing will also be propelled by safe, privacy-protecting real-time analytics. To satisfy the rising need of protecting healthcare data, we predict a need for technical improvements in computing, storage, and communication capacities, which we discuss in this study.

6. REFERENCES

1. Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health affairs, 33(7), 1123-1131.
2. Centers for Medicare and Medicaid Services. (2014). National health expenditure fact sheet. Washington, DC: Centers for Medicare and Medicaid Services. https://www. cms. Gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NHE-Fact-Sheet. html (accessed September 15, 2015).
3. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS quarterly, 1165-1188.
4. Carvalho, R. L. R. D., Campos, C. C., Franco, L. M. D. C., Rocha, A. D. M., & Ercole, F. F. (2017). Incidence and risk factors for surgical site infection in general surgeries 1. Revista latino-americana de enfermagem, 25.
5. De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. Library review, 65(3), 122-135.
6. Gope, P., & Hwang, T. (2015). BSN-Care: A secure IoT-based modern healthcare system using body sensor network. IEEE sensors journal, 16(5), 1368-1376.
7. Hanson, M. A., Powell Jr, H. C., Barth, A. T., Ringgenberg, K., Calhoun, B. H., Aylor, J. H., & Lach, J. (2009). Body area sensor networks: Challenges and opportunities. Computer, 42(1), 58-65.
8. Hill, K. (2012). How target figured out a teen girl was pregnant before her father did. Forbes, Inc, 7, 4-1.
9. Hou, J., Qu, L., & Shi, W. (2019). A survey on internet of things security from data perspectives. Computer Networks, 148, 295-306.
10. Kalejahi, B. K., Meshgini, S., Yariyeva, A., Ndure, D., Maharramov, U., & Farzamnia, A. (2019). Big data security issues and challenges in healthcare. arXiv preprint arXiv:1912.03848.
11. Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. IEEE internet of things journal, 4(5), 1125-1142.
12. Markman, J. (2017). Amazon using AI, Big Data to accelerate profits.
13. Mahmud, R., Koch, F. L., & Buyya, R. (2018, January). Cloud-fog interoperability in IoT-enabled healthcare solutions. In Proceedings of the 19th international conference on distributed computing and networking (pp. 1-10).
14. Marshall, J. A., Trimmer, P. C., Houston, A. I., & McNamara, J. M. (2013). On evolutionary explanations of cognitive biases. Trends in ecology & evolution, 28(8), 469-473.
15. McCann, E. (2013). Kaiser reports second fall data breach. Healthcare IT News, 26.
16. Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. Journal of big data, 2(1), 1-21.

17. Patil, H. K., & Chen, T. M. (2013). Wireless sensor network security. In Computer and Information Security Handbook (pp. 301-322).
18. Paul, A., Pinjari, H., Hong, W. H., Seo, H. C., & Rho, S. (2018). Fog computing-based IoT for health monitoring system. Journal of Sensors, 2018.
19. Ponemon, L. (2012). Third Annual Benchmark Study on Patient Privacy & Data Security. Ponemon Institute, Traverse City, MI (https://www. ponemon. org/news-2/45).
20. Pan, I., Nolan, L. B., Brown, R. R., Khan, R., van der Boor, P., Harris, D. G., & Ghani, R. (2017). Machine learning for social services: a study of prenatal case management in Illinois. American journal of public health, 107(6), 938-944.
21. Petersson, K. M. (2015). Neurobiology of language. In the Catalan Institute for Advanced Studies.
22. Roski, J., Bo-Linn, G. W., & Andrews, T. A. (2014). Creating value in health care through big data: opportunities and policy implications. Health affairs, 33(7), 1115-1122.
23. Toros, H., & Flaming, D. (2017). Prioritizing which homeless people get housing using predictive algorithms. Available at SSRN 2960410.
24. Wang, H., Wang, Z., & Domingo-Ferrer, J. (2018). Anonymous and secure aggregation scheme in fog-based public cloud computing. Future Generation Computer Systems, 78, 712-719.
25. Yang, K., Han, Q., Li, H., Zheng, K., Su, Z., & Shen, X. (2016). An efficient and fine-grained big data access control scheme with privacy-preserving policy. IEEE Internet of Things Journal, 4(2), 563-571.
26. Yang, Y., Wu, L., Yin, G., Li, L., & Zhao, H. (2017). A survey on security and privacy issues in Internet-of-Things. IEEE Internet of things Journal, 4(5), 1250-1258.